Supporting Information

"*In-situ* Cross-Docking" to Simultaneously Address Multiple Targets

Christoph A. Sotriffer, Ingo Dramburg

Protein setup and grid calculations. Structures of the Fab fragment of antisteroid antibody DB3, of the retinol-binding protein (RBP), of chorismate mutase (CM), thrombin (Th), acetylcholinesterase (AChE), and HIV-protease (HPr) were obtained from the PDB (codes 1dbb, 1fen, 2cht, 1ets, 1eve, and 1hvr, respectively). Water molecules and ligands were removed from the files, and polar hydrogens were added to the proteins using the PROTONATE utility distributed with AMBER.[1] Partial atomic charges from the AMBER united atom force field were assigned to the protein atoms, and solvation parameters were added with the ADDSOL utility of AutoDock3.0.[2]

In the actual docking process, AutoDock uses a grid-based representation of the protein. Accordingly, for each point in a regularly spaced grid and for any probe of interest (i.e, any ligand atom types and a probe charge), the interaction energy of that probe with the entire protein is calculated and saved in files that serve as look-up tables for faster energy evaluation upon docking. Accordingly, grids were generated with the help of AutoGrid, using a grid spacing of 1 Å in all cases. First, a separate standard binding-site grid was calculated for each protein (i.e, one for DB3, one for RBP, and so on). Each grid was centered on the corresponding active site and had a dimension of 30 Å × 30 Å × 30 Å, which is sufficiently large to include the entire binding site and significant portions of the surrounding surface. Conventional docking and conventional cross-docking to each single protein were carried out with these grids.

To perform *in-situ* cross-docking, the grids were joined to a single large grid readable by AutoDock. This was done with the three grids of DB3, RBP, and CM, with the three grids of Th, AChE, and HPr, as well as with the grids of all six proteins. Since only already available grid files hat to be manipulated, AutoGrid was not required at this stage. The manipulations corresponded to a linear alignment of the three grids along the x-axis. To avoid docking results across the border of two adjacent grids (which, in structural terms, would be pure artifacts), a "spacer" was inserted between the grids (3 Å in case of three joined grids, 4 Å in case of six joined grids). To each grid point in the spacer region an energy value of +100 kcal/mol was assigned, leading to a "repulsive layer" in which no docking run would terminate.

Ligand setup. Coordinates of progesterone, axerophthene, the endo-oxabicyclic transition state analog binding to CM, NAPAP, Aricept, and XK-203 were taken from the PDB files (1dbb, 1fen, 2cht, 1ets, 1eve, 1hvr) and retrieved in mol2-format from Relibase+.[3] Using Sybyl[4], hydrogens were added to the ligand structures, and atomic charges were assigned according to the Gasteiger-Marsili formalism,[5] which is the type of charges used in calibrating the AutoDock free energy function.[2] Finally, the compounds were setup for docking with the help of AutoTors, the main purpose of which is to define the torsional degrees of freedom to be considered during the docking process. Accordingly, for flexbile docking runs, the following numbers of rotatable bonds were defined: 1 in progesterone, 4 in axerophtene, 3 in the CM transition state analog, 8 in NAPAP, 6 in Aricept, and 10 in XK-203.

Docking. Docking was carried out with AutoDock 3.0[2], using the empirical free energy function and the Lamarckian Genetic algorithm (LGA). The standard protocol used an initial population of 50 or 100 randomly placed individuals, a maximum number of 1.5×10^6 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. For the local search, the pseudo-Solis and Wetts algorithm was applied, using a maximum of 300 iterations per local search. The probability of performing local search was 0.06, and the maximum number of consecutive successes or failures before doubling or halving the local search step size was 4. 10 independent docking runs were carried out for each system and the top-ranked results were analyzed with respect to their rmsd from the experimentally observed position in the crystal structure.

For identification of the fastest successful protocols, variants of the described protocol were used that differed in the maximum number of energy evaluations (ee), as described in the paper. For ee settings larger than the standard 1.5×10^6, variants with larger population sizes (200, 400, and 1000) at a given ee were tested as well and led to the successful protocols for flexible *in-situ* cross-docking of XK-203 to three and six proteins (population size = 1000) and of Aricept to six proteins (population size = 200).

References

1) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 6*; University of California, San Francisco, 1999.

2) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639-1662.

3) Hendlich, M.; Bergner, A.; Günther, J.; Klebe, G. *J Mol Biol* **2003**, *326*, 607-620.

4) *SYBYL Molecular Modeling Software*; Tripos Inc.: St. Louis, MO.

5) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219-3228.